International Journal of Antimicrobial Agents 54 (2019) 265–266

Contents lists available at ScienceDirect

International Journal of Antimicrobial Agents

journal homepage: www.elsevier.com/locate/ijantimicag

Letter to the Editor

If extraordinary data are not first corroborated, we risk being led astray: Occam's razor does not support the existence of plasmid–prophage chimeras. Comment on 'Effects of cryptic prophage regions in a plasmid carrying a carbapenemase gene on survival against antibiotic stress'

Editor: Professor Jian Li

Sir,

The study by Kim and Ko [1], based on next-generation sequencing (NGS) data, presents a very interesting discovery, if confirmed, that a conjugative plasmid has acquired a contiguous locus consisting of multiple non-contiguous segments of prophage DNA, and that each segment contributes to host survival against stress. NGS technology is a powerful new tool aiding microbiological research that should be applied with knowledge of its limitations, especially in the output of the black boxes that are assembly algorithms. Misassembly, which I surmise has occurred here, is one of the most harmful of errors [2].

Extraordinary results, in particular, require corroboration by independent methods in order to have confidence in the data. Unfortunately, I believe that this study falls short for the following reasons.

The authors identified a self-transmissible bla_{NDM-1}-encoding plasmid (pNDM-A1) in a *Klebsiella pneumoniae* isolate by conjugally transferring it to *Escherichia coli* DH5α, a laboratory *E. coli* K12 strain. Using Roche 454 technology, they assembled a plasmid genome, although details of the library preparation and the assembly program/algorithm are not noted. They report an initially incomplete assembly (four contigs of unreported size) with unspecified gaps being closed by PCR and Sanger sequencing, producing a 67 554-bp circular plasmid. The assembled plasmid appears highly related to other enteric plasmids, but additionally is reported to contain a single 16 795-bp locus that shows near-identity to four separate prophage loci (from laboratory strains of *E. coli* K12, namely Rac, e14 and two regions of KplE2, although KpLE2 is no longer considered prophage-related [3]). This sequence assembly is accepted as proof of the plasmid genome, although no physical/genetic characterisation is shown, and there is no indication that any attempt to verify the assembly was done.

Instead, the results were accepted verbatim, and they proceeded to individually delete these four prophage sequences 'from the plasmid' in the DH5α[pNDM-A1] strain using marker-assisted recombineering. No data are presented confirming that the plasmids lost these sequences, or that the kanamycin resistance (Km^R) cassette inserted into the plasmid. The study does not acknowledge the confounding problem that DH5α, in which all these genetic manipulations and assays were done, contains all these prophage-like sequences in its genome [4]. Most likely these deletions actually occurred in the DH5α chromosome, a conclusion supported by their own data. First, the *rac*-deletion primer pair (their Supplementary Table S2, used to amplify the Km^R cassette for recombineering) could not have inserted the cassette into the plasmid since both primers' *rac* sequences are outside of the *rac* sequence in the assembly. Thus, neither end of the marker cassette amplicon has any homology to the plasmid sequence and could not integrate into the plasmid; it is, however, competent to recombine with and delete the region in the native *rac* locus in the DH5α chromosome. Second, the supplemental quantitative real-time PCR data (their Supplementary Fig. S2), purportedly confirming loss of the prophage sequences in the plasmid 'deletions', actually show, in each instance, complete loss of selected prophage gene expression. This is only possible if BOTH the reported prophage-containing plasmid AND the DH5α host simultaneously acquired the Km^R marker and prophage deletion. More likely, the Km^R recombinants acquired the marker/deletions at the chromosomal prophage loci, with the plasmids having no prophage genes; only this scenario is consistent with complete loss of signal. Surprisingly, the authors do not report attempts to produce plasmid deletions that span more than one prophage region (or indeed all four prophage sequences). However, these would be (near) impossible to create in a single step in the DH5α background/genome using recombineering, if, as I hypothesise, the plasmid does not actually have these four non-contiguous *E. coli* prophage regions.

Further analysis of the plasmid sequence also supports a misassembly hypothesis. First, immediately 5′ of the prophage sequences (39–43 kb) is an unusual 4-kb duplication of another plasmid region (3.2–7.2 kb), but divided into a 1780-bp inverted repeat and a 2278-bp direct repeat. Second, a BLAST search using the 17-kb prophage sequence finds all four loci together at >99% identity only in *E. coli* K12 laboratory strains. Third, 1388 bp immediately 3′ of the prophage sequences shows 100% identity to partial (and unlinked) IS*5* and IS*10* sequences from the DH5α genome, highly suggestive of them also originating from the DH5α chromosome. In the absence of any direct evidence that the plasmid–prophage chimera assembly exists other than in silico, the simplest explanation (Occam's razor) fitting all the data is that the unnamed program's assembly algorithm produced a spurious chimera of plasmid contigs and contigs arising from contaminating *E. coli* chromosomal DNA. It then follows that the plasmid–prophage 'deletions' were actually created in the DH5α chromosome, and all deletion strains carry identical pNDM-A1 plasmids, but in host genetic backgrounds deleted for individual prophage loci. This renders any analysis of stress responses irrelevant in reference to the plasmid, and thus negates any conclusions drawn by this study.

DOI of original article: 10.1016/j.ijantimicag.2018.09.002

https://doi.org/10.1016/j.ijantimicag.2018.12.008

Funding

None.

Competing interests

None declared.

Ethical approval

Not required.

References

[1] Kim SY, Ko KS. Effects of prophage regions in a plasmid carrying a carbapenemase gene on survival against antibiotic stress. Int J Antimicrob Agents 2019;53:89–94. doi:10.1016/j.ijantimicag.2018.09.002.
[2] Li M, Wu B, Yan X, Luo J, Pan Y, Wu F-X, et al. PECC: correcting contigs based on paired-end read distribution. Comput Biol Chem 2017;69:178–84. doi:10.1016/j.compbiolchem.2017.03.012.
[3] EcoCyc. *Escherichia coli* K-12 substr. MG1655 Site: PR-Y synonym KpLE2, https://biocyc.org/ECOLI/NEW-IMAGE?type=EXTRAGENIC-SITE&object=PROPH-12; 2017 Accessed 28 October 2018.
[4] Anton BP, Raleigh EA. Complete genome sequence of NEB 5-alpha, a derivative of *Escherichia coli* K-12 DH5α. Genome Announc 2016;4 e01245-16. doi:10.1128/genomeA.01245-16.

Michael G. Jobling
26488 E. Walker Drive, Aurora, CO 80016, USA
E-mail address: michael.jobling@ucdenver.edu

Received 2 November 2018
Accepted 15 December 2018